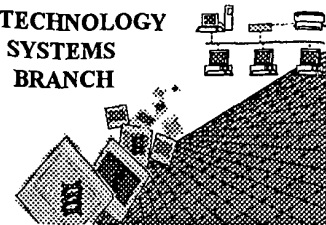


2/15/03

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



**RAW SEQUENCE LISTING**  
**ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/607,077  
Source: OIPG  
Date Processed by STIC: 8/4/2003

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

**<http://www.uspto.gov/web/offices/pac/checker>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 04/24/2003

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 10/607,077

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos  
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length  
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering  
The numbering under each 3<sup>rd</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII  
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length  
Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"  
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9      Use of n's or Xaa's  
    (NEW RULES)  
Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response  
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or Artificial Sequence
- 11      Use of <220>  
Sequence(s)      missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"  
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n  
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/607,077

DATE: 08/04/2003

TIME: 08:36:37

Input Set : A:\ASHBY.1.DIV.ST25.txt

Output Set: N:\CRF4\08042003\J607077.raw

3 <110> APPLICANT: Ashby, Matthew  
 5 <120> TITLE OF INVENTION: Methods for the Survey and Genetic Analysis of Populations  
 7 <130> FILE REFERENCE: ASHBY/1 DIV  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/607,077  
 C--> 9 <141> CURRENT FILING DATE: 2003-06-25  
 9 <150> PRIOR APPLICATION NUMBER: US 09/829855  
 10 <151> PRIOR FILING DATE: 2001-04-10  
 12 <150> PRIOR APPLICATION NUMBER: PCT/US01/11609  
 13 <151> PRIOR FILING DATE: 2001-04-10  
 15 <150> PRIOR APPLICATION NUMBER: US 60/196063  
 16 <151> PRIOR FILING DATE: 2000-04-10  
 18 <150> PRIOR APPLICATION NUMBER: US 60/196258  
 19 <151> PRIOR FILING DATE: 2000-04-11  
 21 <160> NUMBER OF SEQ ID NOS: 244  
 23 <170> SOFTWARE: PatentIn version 3.1  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 16  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: unidentified soil organism  
 30 <400> SEQUENCE: 1  
 31 acgatgagca ctact  
 34 <210> SEQ ID NO: 2  
 35 <211> LENGTH: 16  
 36 <212> TYPE: DNA  
 37 <213> ORGANISM: unidentified soil organism  
 39 <400> SEQUENCE: 2  
 40 acgatgagta ctact  
 43 <210> SEQ ID NO: 3  
 44 <211> LENGTH: 16  
 45 <212> TYPE: DNA  
 46 <213> ORGANISM: unidentified soil organism  
 48 <400> SEQUENCE: 3  
 49 acgatgatga ctact  
 52 <210> SEQ ID NO: 4  
 53 <211> LENGTH: 16  
 54 <212> TYPE: DNA  
 55 <213> ORGANISM: unidentified soil organism  
 57 <400> SEQUENCE: 4  
 58 acgatggatg ctact  
 61 <210> SEQ ID NO: 5  
 62 <211> LENGTH: 16  
 63 <212> TYPE: DNA  
 64 <213> ORGANISM: unidentified soil organism

pp 1-6  
 Does Not Comply  
 Corrected Diskette Needed

invalid <2137> response

see item 10 on Error Summary sheet

16

16

16

16

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/607,077

DATE: 08/04/2003

TIME: 08:36:37

Input Set : A:\ASHBY.1.DIV.ST25.txt

Output Set: N:\CRF4\08042003\J607077.raw

66 <400> SEQUENCE: 5  
67 atgctagtct ggagct 16  
70 <210> SEQ ID NO: 6  
71 <211> LENGTH: 16  
72 <212> TYPE: DNA  
73 <213> ORGANISM: unidentified soil organism  
75 <400> SEQUENCE: 6  
76 atggctgtcg tcagct 16  
79 <210> SEQ ID NO: 7  
80 <211> LENGTH: 16  
81 <212> TYPE: DNA  
82 <213> ORGANISM: unidentified soil organism  
84 <400> SEQUENCE: 7  
85 atggttgtcg tcagct 16  
88 <210> SEQ ID NO: 8  
89 <211> LENGTH: 16  
90 <212> TYPE: DNA  
91 <213> ORGANISM: unidentified soil organism  
93 <400> SEQUENCE: 8  
94 attccgtgcc gtagct 16  
97 <210> SEQ ID NO: 9  
98 <211> LENGTH: 16  
99 <212> TYPE: DNA  
100 <213> ORGANISM: unidentified soil organism  
102 <400> SEQUENCE: 9  
103 cactagtggc gcagct 16  
106 <210> SEQ ID NO: 10  
107 <211> LENGTH: 16  
108 <212> TYPE: DNA  
109 <213> ORGANISM: unidentified soil organism  
111 <400> SEQUENCE: 10  
112 cccccgtgcc gaagct 16  
115 <210> SEQ ID NO: 11  
116 <211> LENGTH: 16  
117 <212> TYPE: DNA  
118 <213> ORGANISM: unidentified soil organism  
120 <400> SEQUENCE: 11  
121 cccccgtgcc gcagct 16  
124 <210> SEQ ID NO: 12  
125 <211> LENGTH: 16  
126 <212> TYPE: DNA  
127 <213> ORGANISM: unidentified soil organism  
129 <400> SEQUENCE: 12  
130 ccccttctct ccagct 16  
133 <210> SEQ ID NO: 13  
134 <211> LENGTH: 16  
135 <212> TYPE: DNA  
136 <213> ORGANISM: unidentified soil organism  
138 <400> SEQUENCE: 13

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/607,077

DATE: 08/04/2003

TIME: 08:36:37

Input Set : A:\ASHBY.1.DIV.ST25.txt

Output Set: N:\CRF4\08042003\J607077.raw

139 ccccggtgcc gcagct	16
142 <210> SEQ ID NO: 14	
143 <211> LENGTH: 16	
144 <212> TYPE: DNA	
145 <213> ORGANISM: unidentified soil organism	
147 <400> SEQUENCE: 14	
148 ccgggtagtc ccagct	16
151 <210> SEQ ID NO: 15	
152 <211> LENGTH: 16	
153 <212> TYPE: DNA	
154 <213> ORGANISM: unidentified soil organism	
156 <400> SEQUENCE: 15	
157 cctccgtgcc gaagct	16
160 <210> SEQ ID NO: 16	
161 <211> LENGTH: 16	
162 <212> TYPE: DNA	
163 <213> ORGANISM: unidentified soil organism	
165 <400> SEQUENCE: 16	
166 cctccgtgcc gcagct	16
169 <210> SEQ ID NO: 17	
170 <211> LENGTH: 16	
171 <212> TYPE: DNA	
172 <213> ORGANISM: unidentified soil organism	
174 <400> SEQUENCE: 17	
175 cctccgtgct gcagct	16
178 <210> SEQ ID NO: 18	
179 <211> LENGTH: 16	
180 <212> TYPE: DNA	
181 <213> ORGANISM: unidentified soil organism	
183 <400> SEQUENCE: 18	
184 cctcggcgcc gcagct	16
187 <210> SEQ ID NO: 19	
188 <211> LENGTH: 16	
189 <212> TYPE: DNA	
190 <213> ORGANISM: unidentified soil organism	
192 <400> SEQUENCE: 19	
193 cctcggtgcc gcagct	16
196 <210> SEQ ID NO: 20	
197 <211> LENGTH: 16	
198 <212> TYPE: DNA	
199 <213> ORGANISM: unidentified soil organism	
201 <400> SEQUENCE: 20	
202 cctcggtgct gcagct	16
205 <210> SEQ ID NO: 21	
206 <211> LENGTH: 16	
207 <212> TYPE: DNA	
208 <213> ORGANISM: unidentified soil organism	
210 <400> SEQUENCE: 21	
211 cctgggtgcc gcagct	16

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/607,077

DATE: 08/04/2003

TIME: 08:36:37

Input Set : A:\ASHBY.1.DIV.ST25.txt

Output Set: N:\CRF4\08042003\J607077.raw

```

214 <210> SEQ ID NO: 22
215 <211> LENGTH: 16
216 <212> TYPE: DNA
217 <213> ORGANISM: unidentified soil organism
219 <400> SEQUENCE: 22
220 cctgtgtgac gaagct 16
223 <210> SEQ ID NO: 23
224 <211> LENGTH: 16
225 <212> TYPE: DNA
226 <213> ORGANISM: unidentified soil organism
228 <400> SEQUENCE: 23
229 ccttggtaac gaagct 16
232 <210> SEQ ID NO: 24
233 <211> LENGTH: 16
234 <212> TYPE: DNA
235 <213> ORGANISM: unidentified soil organism
237 <400> SEQUENCE: 24
238 ccttggtacc gaagct 16
241 <210> SEQ ID NO: 25
242 <211> LENGTH: 16
243 <212> TYPE: DNA
244 <213> ORGANISM: unidentified soil organism
246 <400> SEQUENCE: 25
247 cgccagtgcc gtagct 16
250 <210> SEQ ID NO: 26
251 <211> LENGTH: 16
252 <212> TYPE: DNA
253 <213> ORGANISM: unidentified soil organism
255 <400> SEQUENCE: 26
256 cgccgtgtgcc gtagct 16
259 <210> SEQ ID NO: 27
260 <211> LENGTH: 16
261 <212> TYPE: DNA
262 <213> ORGANISM: unidentified soil organism
264 <400> SEQUENCE: 27
265 cgtccgtgcc gaagct 16
268 <210> SEQ ID NO: 28
269 <211> LENGTH: 16
270 <212> TYPE: DNA
271 <213> ORGANISM: unidentified soil organism
273 <400> SEQUENCE: 28
274 cgtccgtgcc gcagct 16
277 <210> SEQ ID NO: 29
278 <211> LENGTH: 16
279 <212> TYPE: DNA
280 <213> ORGANISM: unidentified soil organism
282 <400> SEQUENCE: 29
283 cgtccgtgcc gcagct 16
286 <210> SEQ ID NO: 30

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/607,077

DATE: 08/04/2003

TIME: 08:36:37

Input Set : A:\ASHBY.1.DIV.ST25.txt

Output Set: N:\CRF4\08042003\J607077.raw

287 <211> LENGTH: 16  
 288 <212> TYPE: DNA  
 289 <213> ORGANISM: unidentified soil organism  
 291 <400> SEQUENCE: 30  
 292 ctcccgtgcc gcagct 16  
 295 <210> SEQ ID NO: 31  
 296 <211> LENGTH: 16  
 297 <212> TYPE: DNA  
 298 <213> ORGANISM: unidentified soil organism  
 300 <400> SEQUENCE: 31  
 301 ctcccgtgcc ggagct 16  
 304 <210> SEQ ID NO: 32  
 305 <211> LENGTH: 16  
 306 <212> TYPE: DNA  
 307 <213> ORGANISM: unidentified soil organism  
 309 <400> SEQUENCE: 32  
 310 ctccggtgcc gcagct 16  
 313 <210> SEQ ID NO: 33  
 314 <211> LENGTH: 16  
 315 <212> TYPE: DNA  
 316 <213> ORGANISM: unidentified soil organism  
 318 <400> SEQUENCE: 33  
 319 ctccgtgcc gaagct 16  
 322 <210> SEQ ID NO: 34  
 323 <211> LENGTH: 16  
 324 <212> TYPE: DNA  
 325 <213> ORGANISM: unidentified soil organism  
 327 <400> SEQUENCE: 34  
 328 ctccgtgcc gcagct 16  
 331 <210> SEQ ID NO: 35  
 332 <211> LENGTH: 16  
 333 <212> TYPE: DNA  
 334 <213> ORGANISM: unidentified soil organism  
 336 <400> SEQUENCE: 35  
 337 ctgccgtgcc gaagct 16  
 340 <210> SEQ ID NO: 36  
 341 <211> LENGTH: 16  
 342 <212> TYPE: DNA  
 343 <213> ORGANISM: unidentified soil organism  
 345 <400> SEQUENCE: 36  
 346 ctgctgtgcc gaagct 16  
 349 <210> SEQ ID NO: 37  
 350 <211> LENGTH: 16  
 351 <212> TYPE: DNA  
 352 <213> ORGANISM: unidentified soil organism  
 354 <400> SEQUENCE: 37  
 355 ctgtcgtgcc gaagct 16  
 358 <210> SEQ ID NO: 38  
 359 <211> LENGTH: 16

↓  
 The types of errors shown exist throughout  
 the Sequence Listing. Please check subsequent  
 sequences for similar errors.

## RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 08/04/2003

PATENT APPLICATION: US/10/607,077

TIME: 08:36:38

Input Set : A:\ASHBY.1.DIV.ST25.txt

Output Set: N:\CRF4\08042003\J607077.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:155; N Pos. 22

Seq#:156; N Pos. 25

Seq#:157; N Pos. 14,15,16,17,18,19,20,21,22

Seq#:171; N Pos. 11

Seq#:233; Xaa Pos. 8



**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/607,077

DATE: 08/04/2003

TIME: 08:36:38

Input Set : A:\ASHBY.1.DIV.ST25.txt

Output Set: N:\CRF4\08042003\J607077.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No  
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:1567 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155 after pos.:0  
L:1585 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156 after pos.:0  
L:1603 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157 after pos.:0  
L:1735 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:171 after pos.:0  
L:2380 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:233 after pos.:0